

## データベース検索音声対話システムにおける履歴を考慮した検索条件の管理

Management of Query Conditions using Dialogue History in Spoken Dialogue Systems

神田 直之 駒谷 和範  
Naoyuki Kanda Kazunori Komatani

尾形 哲也 奥乃 博  
Tetsuya Ogata Hiroshi G. Okuno

## 1 はじめに

音声対話システムの有望なタスクの一つに、電話によるデータベース検索がある。レストランや観光情報などのデータベースに対して、条件を指定して検索を行い、さらに検索結果(特定のレストランなど)に対する詳細情報を聞き出すというタスクである。ユーザは音声で要求をシステムに伝え、得られる検索結果に応じて検索条件の追加・削除を行うことにより、目的の情報を引き出す。ユーザは、自分の好みに最も近いものを試行錯誤しながら探すため、ユーザのゴールは一対話の間にも変化する。したがって、システムはユーザの意図の変化に追従できる必要がある。また、受理できる語彙や対話の展開に自由度を持たせる必要も生じる。

ユーザに自由な発話を許す場合には、対話の進行はユーザの好み・判断に依存するため、これを事前にすべて記述しておくのは不可能である。したがって、音声応答システムで使われる VoiceXML のように、ユーザの次発話に対する認識文法を用意することで音声認識誤りを減らしたり、ユーザ発話による状態の遷移を記述しておくことはできない。つまり、(1) 音声認識誤りへの対処 (2) ユーザの意図に追従可能な対話管理の 2 点が必要となる。

我々は、データベース検索タスクにおける対話の構造に着目し、これらの問題への対処を図る。まずユーザ発話を「検索条件の指定」「情報の提示要求」の 2 つに大別し、これに対応して対話の大局的なモードを規定する。さらに、これらのモードそれぞれに対して、検索条件の動的木を作成し、入力された情報の履歴を表す木構造を保持する。これらは、個別のドメインの知識には依存せず、データベース検索タスクでは一般に成り立つ制約である。

我々はこのモデルに基づき、音声認識誤りを検出するための文脈的信頼度について以前に報告した [1]。本稿では、検索条件の動的木を用いて、ユーザの意図に追従するための対話管理について述べる。音声対話システムでは、マウスやキーボードなど、誤りなく動作を指定できるインタフェースに比べ、ユーザが項目を削除したり状態を変化させる手間が大きい。したがって、ユーザの意図の変化に応じて、単純にスロットを上書きするだけでなく、履歴をもとに適切に検索条件の管理を行う必要がある。

同様のタスクにおいて、[2] ではスロット間の関係を事前に記述することで、検索条件の管理を行っている。本研究では、ドメインごとに知識を手手で記述することなく検索条件の管理を行う。

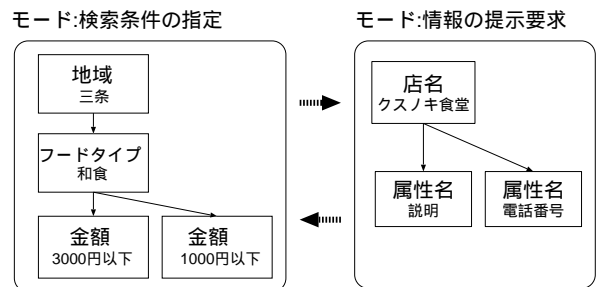


図 1: データベース検索対話における静的な構造と動的な構造

表 1: 各モードにおけるスロットと値の対応

モード	スロット	値
検索条件の指定	地域	三条, ...
	金額	3000 円, ...
	...	...
情報の提示要求	店名	ランゼン, 吉田屋, ...
	属性名	説明, 予算, ...

## 2 データベース検索タスクにおける対話のモデル化

本研究ではタスクドメインとしてレストラン検索を対象とする。データベース検索タスクにおける典型的な対話として、まず希望する検索条件を追加・削除することにより数件まで店を絞り込み、その後絞り込んだ店の具体的な情報に関して質問を行うという手順を想定する。これは、検索条件を指定して店名(関係データベースにおけるキー属性)を絞り込み、絞り込んだ店名が持つ属性に関して情報を得るという過程に相当する。ここでの前者を「検索条件の指定」モード、後者を「情報の提示要求」モードとする(図 1)。

本研究ではこの 2 モード間の遷移を、静的な構造による文脈的制約として規定する。例えば、情報の提示要求モードに移る場合には、検索条件の指定モードで既に絞り込まれた店名に言及するのが自然である。また、情報の提示要求モードで、ある店に関する詳細を聞いている際に、全く別の検索条件が現れるのは音声認識誤りである可能性が高い。このような制約を静的構造による制約とし、音声認識誤りの除去に利用する [1]。各発話は 2 つのモードのいずれかに属し、表 1 に示すスロットに対して値の代入を行うものとする。各モードで現れる表現や語彙は異なるため、各モードごとに用意した想定質問文とのマッチング [3] の

<sup>0</sup> 京都大学 大学院 情報学研究科 知能情報学専攻

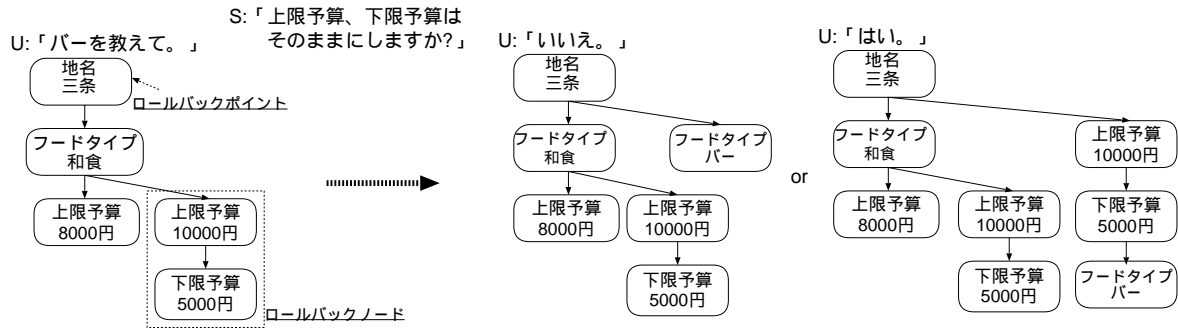


図 2: ロールバック時の動的木の変化の例

結果などから、ある発話がどのモードであるかを推定する。さらに対話の動的な構造として、対話の履歴を表す木構造を定義する。木の各ノードはスロットと値の組で構成されており、上記の 2 モードそれぞれに対して木を保持する(図 1)。発話によりスロットに値が追加されればノードが追加され、値が削除されるとノードも削除される。対話のある時点での検索条件は、木のルートから最も右の子を辿っていくことで得られる(以下これらのノードを参照ノードと呼ぶ)。木構造は具体的には以下の規則により作成される。

- 新たに追加するノードと同じスロットを持つノードが参照ノード中になければ、参照ノードの子の位置にノードを追加する
- 同じスロットを持つノードが参照ノード中にある場合、そのノードの弟の位置にノードを追加する
- 削除したいノードの下に部分木が存在しない場合、そのノードを削除する。
- 削除したいノードの下に部分木が存在する場合、削除するノードの弟の位置にその部分木を移動する

### 3 ユーザの意図の変化に追従した検索条件の管理

データベース検索音声対話システムでは、検索結果に応じてユーザが検索条件を追加・削除する。このとき、ユーザがすでに入力済みの条件を上書きする場合には、ユーザの意図に変化が生じたこととみなし、その変化の影響が及ばないポイントに戻って対話を再開するのが望ましい。このポイントを本研究ではロールバックポイントとする。システムが適切な応答を行うためには、このロールバックポイント、すなわち対話をどこまで戻すかを決定する必要がある。

本研究では、動的構造を利用してロールバックポイントを決める。ユーザの発話があった場合には、その内容が動的構造に新規ノードとして追加される。このとき、動的木の生成規則に基づき、同じスロットを持つノードが参照ノード中にある場合、そのノードの弟の位置に新規ノードは追加される。ここで、追加されるノードの親ノードをロールバックポイントとする。また、参照ノードのうち、新規ノードとは異なるスロットを持つノードを、ロールバックノードと呼ぶ。

本研究では、ロールバックノードが 2 個以上あった場合に、ロールバックノードを削除せずに残すかどうかをユー

ザに確認するという対話戦略を採る。確認に対してユーザが肯定の回答をした場合は、ユーザは新規ノード中の条件のみを入れ替えたいという意図を持つ場合に相当するため、ロールバックポイントと新規ノードとの間に、ロールバックノードを挿入する<sup>1</sup>。確認に対しユーザが否定した場合には、ユーザの意図がロールバックポイントまで戻ったとみなせるため、新規ノードの追加だけを行い、対話を続ける。

図 2 に、ユーザの意図が変わったときの動的木の例を示す。最も左の状態に対して、新たに【フードタイプ: パー】が発話された場合の動作を表す。このとき【地名: 三条】がロールバックポイント【上限予算: 10000 円】【下限予算: 5000 円】がロールバックノードとなる。ロールバックノードを残すかというシステムからの質問に対して、ユーザが否定の回答をした場合には、ユーザの意図がロールバックポイントまで戻ったものとして、ロールバックポイントの子に新規ノード(【フードタイプ: パー】)のみを追加し、以降の対話を進める。ユーザが肯定の回答をした場合には、ロールバックポイントの新たな子として、ロールバックノードと新規ノードを追加し、新たな動的木とする。

### 4 おわりに

本稿では、データベース検索音声対話のモデルについて述べ、その動的構造から、検索条件のロールバックを行う範囲を決定する方法について述べた。これにより、単純に検索条件を上書きするだけでなく、対話の履歴に基づいて検索条件を管理をすることができる。今後被験者を使った評価実験を行い、定量的な評価を報告する予定である。なお本研究の一部は、科研費、21 世紀 COE の支援を受けた。

### 参考文献

- [1] 神田直之, 駒谷和範, 尾形哲也, 奥乃博. 音声対話システムにおける話題の構造を用いた効率的な対話管理. 情報処理学会第 66 回全国大会講演論文集, 4T-5, 2004.
- [2] D. Goddeau, H. Meng, J. Polifroni, S. Seneff, and S. Busayapongchai. A form-based dialogue manager for spoken language applications. In *Proc. ICSLP*, 1996.
- [3] 駒谷和範, 河原達也, 清田陽司, 黒橋禎夫, Pascale Fung. 柔軟な言語モデルとマッチングを用いた音声によるレストラン検索システム. 情報処理学会研究報告, 2001-SLP-39-30, 2001.

<sup>1</sup>入力された時期がより新しいノードが子になるという動的木の定義から、新規ノードは、ロールバックノードの子とする。